AN ABRAMOV FORMULA FOR STATIONARY SPACES OF DISCRETE GROUPS

YAIR HARTMAN, YURI LIMA, AND OMER TAMUZ

Abstract. Let (G, μ) be a discrete group equipped with a generating probability measure, and let Γ be a finite index subgroup of G. A μ-random walk on G, starting from the identity, returns to Γ with probability one. Let θ be the hitting measure, or the distribution of the position in which the random walk first hits Γ.

We prove that the Furstenberg entropy of a (G, μ)-stationary space, with respect to the induced action of (Γ, θ), is equal to the Furstenberg entropy with respect to the action of (G, μ), times the index of Γ in G. The index is shown to be equal to the expected return time to Γ.

As a corollary, when applied to the Furstenberg-Poisson boundary of (G, μ), we prove that the random walk entropy of (Γ, θ) is equal to the random walk entropy of (G, μ), times the index of Γ in G.

1. Introduction and statement of results

Let G be a countable, discrete group, equipped with a generating probability measure μ, and consider the μ-random walk on G: at each step, the increments are chosen independently according to the measure μ. We study recurrent subgroups: Γ is a recurrent subgroup of the pair (G, μ) if the μ-random walk on G almost surely visits Γ infinitely often.

A natural measure on a recurrent subgroup Γ is the hitting measure θ, which is the distribution of the first element of Γ visited by a μ-random walk on G. Our main result answers a question of Furstenberg regarding the relation between the μ-random walk on G and the corresponding θ-random walk on Γ, from the point of view of their stationary actions.

The main objects of study of this article are stationary spaces. A probability space (X, ν) is (G, μ)-stationary if it is a measurable G-space, and the convolution $\mu * \nu$ is equal to ν. Essentially, the action of G on X leaves ν invariant "on the average", when this average is taken over μ.

A G-space is in particular a Γ-space. Furthermore, it is known that if (X, ν) is (G, μ)-stationary then it is also (Γ, θ)-stationary.

The Furstenberg entropy of a (G, μ)-stationary space (X, ν), denoted by $h_\mu(X, \nu)$, measures the average deformation of ν under the action of G. We prove that the entropies of (X, ν) with respect to the actions of G and of Γ are easily related.

Date: April 25, 2012.

2010 *Mathematics Subject Classification.* Primary: 37A50, 46L55, 60J50. Secondary: 60J05.

Key words and phrases. Abramov formula, Furstenberg entropy, random walk, random walk entropy, stationary space.

Yuri Lima is supported by the European Research Council, grant 239885. Omer Tamuz is supported by ISF grant 1300/08, and is a recipient of the Google Europe Fellowship in Social Computing, and this research is supported in part by this Google Fellowship.

Theorem 1.1 (Abramov formula for stationary spaces)**.** *Let G be a countable, discrete group with a generating probability measure μ with finite entropy. Let Γ be a finite index subgroup of G, with hitting measure θ. Then, for any (G,μ)-stationary space (X,ν),*

$$h_\theta(X,\nu) = [G:\Gamma]\cdot h_\mu(X,\nu). \tag{1.1}$$

In classical ergodic theory, if $T:(X,\nu)\to(X,\nu)$ is a measure-preserving transformation and $A\subset X$ has positive measure, one can define a measure-preserving induced transformation on A as the first return of $x\in A$ to A, under repeated applications of T. Abramov related in [1] the entropy of T with the entropy of the induced map. More specifically, he proved that the entropy of the induced map is equal to $1/\nu(A)$ times the entropy of the initial map, thus relating the ratio of entropies to the fraction of X occupied by A. If one thinks of the index $[G:\Gamma]$ as the portion that Γ occupies in G, then Theorem 1.1 makes the analogous statement. We therefore call Theorem 1.1 an *Abramov formula*.

An ingredient of the proof of Theorem 1.1 is the following result, which states that the expected return time of a μ-random walk to Γ is equal to its index in G. In particular, the expected return time is independent of the measure μ. This holds in the more general setting of topological groups and open subgroups.

Theorem 1.2 (Kac formula for subgroups)**.** *Let G be a second countable topological group with a generating probability measure μ, and let Γ be an open, recurrent subgroup of (G,μ). Then the expected return time $\mathbb{E}[\tau]$ of the μ-random walk to Γ is equal to the index of Γ in G:*

$$\mathbb{E}[\tau] = [G:\Gamma].$$

We call Theorem 1.2 a *Kac formula* in another analogy with classical ergodic theory: there, Kac's formula states that the expected return time of $x\in A$ to A, under repeated applications of T, is equal to $1/\nu(A)$. We indeed use the classical Kac formula to establish this theorem.

Finally, we apply Theorem 1.1 to the special case that (X,ν) is the Furstenberg-Poisson boundary of (G,μ), which is naturally isomorphic to the Furstenberg-Poisson boundary of (Γ,θ). Using the Kaimanovich-Vershik entropy characterization of the Furstenberg-Poisson boundary [7], we get that $h(\Gamma,\theta)$, the random walk entropy of (Γ,θ), is equal to the index of Γ in G times $h(G,\mu)$, the random walk entropy of (G,μ).

Corollary 1.3. *Let G be a countable, discrete group with a generating probability measure μ with finite entropy. Let Γ be a finite index subgroup of G, with hitting measure θ. Then*

$$h(\Gamma,\theta) = [G:\Gamma]\cdot h(G,\mu). \tag{1.2}$$

Note that by Theorem 1.2 we could write Eq. 1.2 as

$$h(\Gamma,\theta) = \mathbb{E}[\tau]\cdot h(G,\mu).$$

Likewise, Eq. 1.1 can be written as

$$h_\theta(X,\nu) = \mathbb{E}[\tau]\cdot h_\mu(X,\nu).$$

The paper is organized as follows. In Section 2 we introduce the basic notations and definitions as well as the necessary background for the sequel. Section 3 is devoted to the proof of Theorem 1.2, and Section 4 to the proof of Theorem 1.1.

To this matter, Appendix A treats the required results of Markov chains, adapted to our context.

2. Preliminaries

The following definitions are mostly standard. The notation is adapted from Furman [4], who provides an excellent exposition to random walks on groups and Furstenberg-Poisson boundary theory.

2.1. **Random walks on groups and random walk entropy.** Let G be a countable, discrete group with identity e, and let μ be a *generating* probability measure on G, so that the semigroup generated by its support,

$$\mathrm{supp}(\mu) = \{g \in G\,;\, \mu(g) > 0\},$$

is equal to the whole of G. Let $X_1, X_2, \ldots$ be a sequence of G-valued independent random variables each with law μ. The Markov chain $\{Z_n\}_{n=1}^{\infty}$ defined as

$$Z_n = X_1 X_2 \cdots X_n$$

is called the *μ-random walk* on G. Formally, if we let $(\Omega, \mathbb{P}) = (G, \mu)^{\mathbb{N}}$, then

$$\begin{array}{cccc} X_n & : & (\Omega, \mathbb{P}) & \longrightarrow & G \\ & & \omega & \longmapsto & X_n(\omega) = \omega_n, \end{array}$$

and

$$\begin{array}{cccc} Z_n & : & (\Omega, \mathbb{P}) & \longrightarrow & G \\ & & \omega & \longmapsto & Z_n(\omega) = X_1(\omega) \cdots X_n(\omega). \end{array}$$

We will occasionally consider random walks starting from some $g \in G$, in which case $Z_n = g X_1 X_2 \cdots X_n$.

Let μ be a probability measure on G. The standard definition for the *entropy* of μ is

$$H(\mu) = -\sum_{g \in G} \mu(g) \cdot \log \mu(g),$$

where $0 \cdot \log 0 = 0$. Throughout this work, we only consider finite entropy probability measures, i.e., measures for which H is finite. As Kaimanovich and Vershik point out, measures with infinite entropy are somewhat different in nature (see the remark on page 465 in [7]).

Let μ_1 and μ_2 be probability measures on G. The *convolution* of μ_1 and μ_2 is defined by

$$[\mu_1 * \mu_2](g) = \sum_{g' \in G} \mu_1(g') \mu_2(g'^{-1} g).$$

Equivalently, $\mu_1 * \mu_2$ is the push-forward of the product measure $\mu_1 \times \mu_2$, under the product map $(g_1, g_2) \mapsto g_1 g_2$ from $G \times G$ to G. By standard information theoretical inequalities it holds that

$$H(\mu_1 * \mu_2) \leq H(\mu_1) + H(\mu_2).$$

In particular, for a probability measure μ on G, if μ^n denotes the nth convolution of μ with itself, the sequence $H(\mu^n)/n$ converges, and so we can consider the following definition of random walk entropy [2].

Definition 2.1. The *random walk entropy*, also known as *Avez entropy* or *asymptotic entropy*, of the pair (G, μ) is

$$h(G,\mu) = \lim_{n\to\infty} \frac{1}{n} H(\mu^n).$$

The measure μ^n is the distribution of the position of a μ-random walk in its nth step. In a sense, $h(G, \mu)$ measures the rate of escape to infinity of a μ-random walk on G.

In this article, we are interested in the relation between the μ-random walk on G and its induced random walk on a recurrent subgroup.

Definition 2.2. A subgroup Γ of G is *μ-recurrent* if, for $\mathbb{P}$-almost every $\omega \in \Omega$, there exists an $n \geq 1$ ($\Leftrightarrow$ infinitely many values of n) such that $Z_n(\omega) \in \Gamma$.

The recurrence property is equivalent to the existence of a *return time map* $\tau : \Omega \to \mathbb{N}$ defined $\mathbb{P}$-almost everywhere by

$$\tau(\omega) = \min\{n \geq 1\,;\, Z_n(\omega) \in \Gamma\}$$

and thus of a *hitting map*

$$\begin{array}{rccc} \Phi & : & \Omega & \longrightarrow & \Gamma \\ & & \omega & \longmapsto & Z_{\tau(\omega)}(\omega). \end{array}$$

The random variable Φ is the first element of Γ hit by the random walk. A useful related definition is that of avoiding sets A_n.

Definition 2.3. Let Γ be a subgroup of G. The nth *avoiding set* A_n is

$$A_n = \{(g_1, \ldots, g_n) \in G^n\,;\, g_1 \cdots g_k \notin \Gamma \text{ for } k \leq n\}. \tag{2.1}$$

Equivalently, A_n is the set of all length n walks on G that do not hit Γ. Evidently,

$$\tau > n \quad \Leftrightarrow \quad (X_1, \ldots, X_n) \in A_n.$$

Let Γ be a μ-recurrent subgroup of G. The distribution of Φ defines a natural probability measure on Γ, called the hitting measure.

Definition 2.4. Let Γ be a μ-recurrent subgroup of G. The *hitting measure* of Γ is the probability measure θ on Γ defined by

$$\theta(\gamma) = \mathbb{P}\left[\Phi = \gamma\right], \quad \gamma \in \Gamma.$$

Equivalently, θ is equal to the push-forward of $\mathbb{P}$ under Φ. Another description, which will be useful for our purposes, is the following: for each $n \geq 1$ define $\theta^{(n)}(\gamma)$ to be the probability that a μ-random walk on G will hit Γ first at step n, exactly at $\gamma \in \Gamma$:

$$\theta^{(n)}(\gamma) = \mathbb{P}\left[(X_1, \ldots, X_{n-1}) \in A_{n-1} \text{ and } Z_n = \gamma\right].$$

Note that $\theta^{(n)}$ is not necessarily a probability measure, and that

$$\theta = \sum_{n\in\mathbb{N}} \theta^{(n)}.$$

With respect to the hitting time map τ, one can divide recurrent subgroups into two classes.

Definition 2.5. The *expected return time* to Γ is equal to

$$\mathbb{E}[\tau] = \int_\Omega \tau(\omega) d\mathbb{P}(\omega).$$

Γ is called *positive recurrent* if $\mathbb{E}[\tau] < \infty$ and *null recurrent* if $\mathbb{E}[\tau] = \infty$.

2.2. **Stationary spaces and Furstenberg entropy.** Let X be a *measurable G-space*. By this we mean that G acts on X respecting the group operation of G, and that the action map

$$\begin{array}{rcl} G \times X & \longrightarrow & X \\ (g, x) & \longmapsto & gx \end{array}$$

is measurable. Let ν be a probability measure on X. For $g \in G$ we denote by $g\nu$ the probability measure on X defined by

$$g\nu(A) = \nu(g^{-1}A),$$

where A is a measurable subset of X.

The *convolution* of the measures μ on G and ν on X, denoted by $\mu * \nu$, is the probability measure on X defined as the image of the product measure $\mu \times \nu$ under the above action map. Equivalently,

$$\mu * \nu = \sum_{g \in G} \mu(g) \cdot g\nu.$$

Definition 2.6. A probability measure ν on a G-space X is called *μ-stationary* if $\mu * \nu = \nu$. In this case, (X, ν) is called a *(G, μ)-stationary space*.

This can be interpreted as saying that ν is invariant "on the average". Every such stationary action preserves the measure class of ν, that is, $\nu(A) = 0$ if and only if $g\nu(A) = 0$ for every $g \in G$ and measurable $A \subset X$ (see, e.g., Nevo and Zimmer [9]). In this case, the *Radon-Nikodym cocycle* $\rho : G \times X \to \mathbb{R}$ is defined as

$$\rho(g, x) = -\log \frac{dg^{-1}\nu}{d\nu}(x). \tag{2.2}$$

Note that ρ indeed satisfies the additive cocycle relation:

$$\rho(gg_1, x) = \rho(g, g_1x) + \rho(g_1, x). \tag{2.3}$$

Define $\varphi : G \to \mathbb{R}$ by

$$\varphi(g) = \int_X \rho(g, x) d\nu(x) \tag{2.4}$$

which, by Jensen's inequality, is nonnegative and equal to zero if and only if $g^{-1}\nu = \nu$. Unlike the measure-preserving transformations of classical ergodic theory, here each $g \in G$ may deform ν, and this deformation is quantified by $\varphi(g)$. In terms of information theory, $\varphi(g)$ is the *Kullback-Leibler divergence* $D_{KL}(g\nu || \nu)$ between the measures $g\nu$ and ν. The average deformation is measured as follows.

Definition 2.7. The *Furstenberg entropy* of a (G, μ)-stationary space (X, ν) is

$$h_\mu(X, \nu) = \sum_{g \in G} \mu(g) \cdot \varphi(g) = \sum_{g \in G} \mu(g) \cdot \int_X -\log \frac{dg^{-1}\nu}{d\nu} d\nu(x) d\mu(g).$$

The Furstenberg entropy and the random walk entropy relate to each other in the following way.

Theorem 2.8 (Kaimanovich and Vershik [7]). *If (X,ν) is a (G,μ)-stationary space, then*

$$h_\mu(X,\nu) \le h(G,\mu),$$

with equality if (X,ν) is the Furstenberg-Poisson boundary of (G,μ).

The Furstenberg-Poisson boundary of (G,μ) is introduced in the next section.

2.3. **Furstenberg-Poisson boundary.**

Definition 2.9. A function $h : G \to \mathbb{R}$ is *μ-harmonic* if it satisfies the μ-mean value property

$$h(g) = \sum_{g_1 \in G} \mu(g_1) \cdot h(gg_1)$$

for all $g \in G$.

We will call a function *harmonic*, without explicit reference to the measure, whenever the measure is obvious from the context.

Let $\mathcal{H}^\infty(G,\mu)$ be the space of all bounded harmonic functions on G and $L^\infty(X,\nu)$ be the space of bounded functions on X, with respect to the measure class of ν. When (X,ν) is (G,μ)-stationary, to each $f \in L^\infty(X,\nu)$ we associate the bounded harmonic function $F_\mu(f) \in \mathcal{H}^\infty(G,\mu)$ given by

$$[F_\mu(f)](g) = g\nu(f) = \int_X f(gx)d\nu(x)\,, \quad g \in G. \tag{2.5}$$

This defines a linear map F_μ from $L^\infty(X,\nu)$ to $\mathcal{H}^\infty(G,\mu)$, called the *Furstenberg transform*. The next lemma shows that the image of F_μ is indeed in $\mathcal{H}^\infty(G,\mu)$, and that this condition is equivalent to the (G,μ)-stationarity of (X,ν).

Lemma 2.10. *Let (X,ν) be a G-space and the Furstenberg transform map $F_\mu : L^\infty(X,\nu) \to L^\infty(G,\mu)$ defined as in Eq. 2.5. Then $F_\mu(L^\infty(X,\nu)) \subset \mathcal{H}^\infty(G,\mu)$ if and only if (X,ν) is (G,μ)-stationary.*

Proof. First assume that $F_\mu(L^\infty(X,\nu)) \subset \mathcal{H}^\infty(G,\mu)$. Let $f \in L^\infty(X,\nu)$ and $h = F_\mu(f)$. Then

$$[\mu * \nu](f) = \sum_{g \in G} \mu(g) \cdot g\nu(f) = \sum_{g\in G} \mu(g) \cdot h(g) = h(e) = \nu(f),$$

where in the third equality we used the μ-harmonicity of h at e. Because f is arbitrary, it follows that $\mu * \nu = \nu$.

Conversely, if there exists an f_0 such that $h_0 = F_\mu(f_0)$ is not harmonic at some $g_0 \in G$ then, assuming without loss of generality that $g_0 = e$,

$$[\mu * \nu](f_0) = \sum_{g\in G} \mu(g) \cdot g\nu(f_0) = \sum_{g \in G} \mu(g) \cdot h_0(g) \ne h_0(e) = \nu(f_0),$$

and so $\mu * \nu \ne \nu$. □

Among the many characterizations of the Furstenberg-Poisson boundary, we consider the following.

Definition 2.11. The *Furstenberg-Poisson boundary* of (G,μ) is the unique μ-stationary G-space (X,ν) for which the Furstenberg transform F_μ is an isometric bijection from $L^\infty(X,\nu)$ to $\mathcal{H}^\infty(G,\mu)$.

The aforementioned uniqueness is up to isomorphism of (G,μ)-stationary spaces. For precise definitions and constructions, we refer the reader to Bader and Shalom [3] or Furstenberg and Glasner [6].

2.4. **Induced action.** Let (X,ν) be a (G,μ)-stationary space and Γ a μ-recurrent subgroup of G with hitting measure θ. In this section it is shown that the restricted Γ-action on X is θ-stationary. This follows from the stronger fact that there exists a one-to-one correspondence between μ-harmonic functions on G and θ-harmonic functions on Γ, according to Theorem 2.13 below. In particular, (G,μ) and (Γ,θ) have isomorphic Furstenberg-Poisson boundaries. These results are due to Furstenberg [5]. We include the proof for completeness.

If Γ is recurrent then it is also recurrent for a random walk starting from an arbitrary $g \in G$: since μ is generating, the random walk hits g with positive probability, and conditioned on that returns to Γ with probability one. Hence a random walk that starts at g will return to Γ with probability one.

Accordingly, let θ_g denote the hitting measure on Γ of the μ-random walk starting at $g \in G$, and let $\theta_g^{(n)}(\gamma)$ be the probability that a μ-random walk starting at g will hit Γ first at step n, exactly at γ.

Lemma 2.12. *Let $h \in \mathcal{H}^\infty(G,\mu)$. Then*

$$h(g) = \theta_g(h)\,, \quad \textit{for all } g \in G.$$

The intuition behind this lemma is the following: Γ is recurrent, and hence the μ-random walk hits Γ almost surely. Γ can be therefore be viewed as a boundary of the random walk, and so, as in the solution of the classical Dirichlet problem, the value of a harmonic function h at some $g \in G$ is equal to the average value of h on this boundary, weighted according to the hitting measure of a random walk starting at g. That is, $h(g) = \theta_g(h)$. In particular, h is determined by its values on Γ.

Proof. Since h is a bounded harmonic function, the sequence of random variables $M_1, M_2, \dots$ defined by

$$M_n = h(Z_n)$$

is a bounded martingale. Note that τ is a stopping time, and hence, by the optional stopping theorem,

$$h(g) = \mathbb{E}_g\left[h(Z_1)\right] = \mathbb{E}_g\left[h(Z_\tau)\right] = \theta_g(h),$$

where $\mathbb{E}_g\left[\cdot\right]$ denotes expectation for random walks starting at g. $\square$

The next theorem shows that not only is h determined by its values on Γ, but that the restriction of h to Γ is an isometric bijection between μ-harmonic functions on G and θ-harmonic functions on Γ.

Theorem 2.13 (Furstenberg [5])**.** *The restriction map*

$$\begin{array}{cccc} \Psi & : & \mathcal{H}^\infty(G,\mu) & \longrightarrow & \mathcal{H}^\infty(\Gamma,\theta) \\ & & h & \longmapsto & h|_\Gamma \end{array} \tag{2.6}$$

is an isometric bijection between $\mathcal{H}^\infty(G,\mu)$ and $\mathcal{H}^\infty(\Gamma,\theta)$.

Proof. First we show that the image of Ψ is in $\mathcal{H}^\infty(\Gamma,\theta)$: for $h \in \mathcal{H}^\infty(G,\mu)$, we show that $\Psi(h)$ is θ-harmonic. Let $h' = \Psi(h)$, so that $h'(\gamma) = h(\gamma)$ for $\gamma \in \Gamma$. By Lemma 2.12 above, we have that

$$h'(\gamma) = h(\gamma) = \theta_\gamma(h) = \sum_{\lambda\in\Gamma} \theta_\gamma(\lambda)\cdot h(\lambda).$$

Since in the last expression we evaluate h only on Γ then we can replace h with h', so that

$$h'(\gamma) = \sum_{\lambda\in\Gamma} \theta_\gamma(\lambda)\cdot h'(\lambda).$$

We claim that

$$\theta_\gamma(\lambda) = \theta(\gamma^{-1}\lambda). \tag{2.7}$$

To see this, couple two μ-random walks to perform the same increments, where one starts at γ and the other at the identity. Then the first walk visits λ exactly when the second walk visits $\gamma^{-1}\lambda$.

Hence

$$h'(\gamma) = \sum_{\lambda\in\Gamma} \theta(\lambda)\cdot h'(\gamma\lambda),$$

and so $h' = \Psi(h)$ is indeed θ-harmonic.

We next show that Ψ is a bijection. By Lemma 2.12 h is determined by its values on Γ, and therefore Ψ is one-to-one. To show that Ψ is onto, given $h' \in \mathcal{H}^\infty(\Gamma,\theta)$, define $h \in \mathcal{H}^\infty(G,\mu)$ by $h(g) = \theta_g(h')$. Observe that $\Psi(h) = h'$: for $\gamma \in \Gamma$,

$$h(\gamma) = \theta_\gamma(h') = \sum_{\lambda\in\Gamma} \theta_\gamma(\lambda)\cdot h'(\lambda) = h'(\gamma),$$

by Eq. 2.7 and the θ-harmonicity of γ. We now show that h is indeed μ-harmonic.

Express the probability to first hit Γ at γ when starting a μ-random walk at g using the law of conditional probabilities, by conditioning on the first step, and separating the sum to the events that the first step either hit or did not hit Γ:

$$\theta_g(\gamma) = \sum_{gg_1\in\Gamma} \mu(g_1)\cdot\delta_{gg_1}(\gamma) + \sum_{gg_1\notin\Gamma} \mu(g_1)\cdot\theta_{gg_1}(\gamma),$$

where δ_g is the Dirac measure concentrated on g. Hence we can write this equality as an equality of measures:

$$\theta_g = \sum_{gg_1\in\Gamma} \mu(g_1)\cdot\delta_{gg_1} + \sum_{gg_1\notin\Gamma} \mu(g_1)\cdot\theta_{gg_1}.$$

But for $gg_1 \in \Gamma$ it holds that $\delta_{gg_1}(h') = h'(gg_1) = \theta_{gg_1}(h')$. Hence

$$\theta_g(h') = \sum_{g_1\in G} \mu(g_1)\cdot\theta_{gg_1}(h').$$

The left hand side of the equation above is equal to $h(g)$, and the right hand side is the μ-mean value property of h at g, and so h is indeed μ-harmonic.

To conclude the proof we show that Ψ preserves the sup norms. Clearly, $\|h'\| \le \|h\|$. Now if $\|h'\| < \|h\|$, then there exists some $g_0 \in G$ with $|h(g_0)| > |h(\gamma)|$ for all $\gamma \in \Gamma$. But $h(g_0) = \theta_{g_0}(h)$, that is $h(g_0)$ is an average of values of the form $h(\gamma)$, in contradiction. $\square$

Corollary 2.14. *Every (G,μ)-stationary space is also (Γ,θ)-stationary. Furthermore, (G,μ) and (Γ,θ) have the same Furstenberg-Poisson boundary.*

Proof. If (X,ν) is (G,μ)-stationary, then $F_\theta = \Psi \circ F_\mu$ maps $L^\infty(X,\nu)$ to $\mathcal{H}^\infty(\Gamma,\theta)$. By Lemma 2.10, this means that ν is θ-stationary.

If (X,ν) is the Furstenberg-Poisson boundary of (G,μ), then $F_\theta = \Psi \circ F_\mu$ is a composition of isometric bijections and thus an isometric bijection as well. □

3. Expected hitting time and index: a Kac formula

The goal of this section is to prove Theorem 1.2, namely that the expected return time to a recurrent subgroup is equal to its index, whether the subgroup is positive or null recurrent. We do this by inducing a Markov chain on the quotient $\Gamma\backslash G$.

Let Γ be a subgroup of G,

$$\Gamma\backslash G = \{\Gamma g\,;\, g \in G\}$$

the set of right cosets and $\pi : G \to \Gamma\backslash G$ the projection map. G naturally acts on $\Gamma\backslash G$ by right multiplication,

$$\begin{array}{ccc} \Gamma\backslash G \times G & \longrightarrow & \Gamma\backslash G \\ (\Gamma g, g_1) & \longmapsto & \Gamma g g_1\,. \end{array}$$

Since $((\Gamma h)g)g_1 = (\Gamma h)(gg_1)$, each map

$$\begin{array}{ccccc} g & : & \Gamma\backslash G & \longrightarrow & \Gamma\backslash G \\ & & \Gamma g_1 & \longmapsto & (\Gamma g_1)g = \Gamma g_1 g \end{array}$$

is a bijection.

The Markov chain $\{Z_n\}_{n=1}^\infty$ projects under π to a Markov chain $\{Y_n\}_{n=1}^\infty$ which has special properties, according to the following lemma.

Lemma 3.1. *Let G be a second countable topological group with a generating probability measure μ, and let Γ be an open, recurrent subgroup of (G,μ). Then $\{Y_n\}_{n=1}^\infty$ is a time-independent Markov chain on $\Gamma\backslash G$. Furthermore, it is*

(a) doubly stochastic,
(b) irreducible, and
(c) recurrent.

Proof. Assume first that G is discrete. For $\Gamma g_1, \Gamma g_2 \in \Gamma\backslash G$, denote the transition probabilities by

$$p(\Gamma g_1, \Gamma g_2) = \sum_{\substack{g\in G \\ \Gamma g_1 g = \Gamma g_2}} \mu(g).$$

This indeed defines a stochastic matrix $\mathbf{P} = (p(\Gamma g_1, \Gamma g_2))_{\Gamma\backslash G \times \Gamma\backslash G}$ for which

$$\begin{aligned} \mathbb{P}\left[Y_n = \Gamma g_n \,|\, Y_1 = \Gamma g_1, \ldots, Y_{n-1} = \Gamma g_{n-1}\right] &= \mathbb{P}\left[Y_n = \Gamma g_n \,|\, Y_{n-1} = \Gamma g_{n-1}\right] \\ &= p(\Gamma g_{n-1}, \Gamma g_n). \end{aligned}$$

Hence $\{Y_n\}_{n=1}^\infty$ is a time-independent Markov chain on $\Gamma\backslash G$ with transition matrix $\mathbf{P}$. We now prove (a), (b) and (c).

(a) The sum of the column of $\mathbf{P}$ associated with Γg_2 is

$$\sum_{\Gamma g_1 \in \Gamma\backslash G} p(\Gamma g_1, \Gamma g_2) = \sum_{\Gamma g_1 \in \Gamma\backslash G} \sum_{\substack{g\in G \\ \Gamma g_1 g = \Gamma g_2}} \mu(g).$$

Changing the order of summation and rearranging we get

$$\sum_{\Gamma g_1 \in \Gamma\backslash G} p(\Gamma g_1, \Gamma g_2) = \sum_{g\in G} \mu(g) \cdot |\{\Gamma g_1 \in \Gamma\backslash G\,;\, \Gamma g_1 g = \Gamma g_2\}| = \sum_{g\in G} \mu(g) = 1,$$

where in the second equality we used that each $g : \Gamma\backslash G \to \Gamma\backslash G$ is a bijection.

(b) The fact that μ is generating guarantees that $\{Z_n\}_{n=1}^\infty$ is an irreducible Markov chain, and the irreducibility property descends to $\{Y_n\}_{n=1}^\infty$.

(c) Note that Z_n belongs to Γ if and only if $Y_n = \Gamma e$, the trivial coset. Therefore, since $\{Z_n\}_{n=1}^\infty$ is recurrent to Γ, $\{Y_n\}_{n=1}^\infty$ is recurrent to the trivial coset Γe.

Consider now the general case that Γ is an open, recurrent subgroup of a second countable group G. Then the quotient space $\Gamma\backslash G$ is countable, and so we can define the Markov chain $\{Y_n\}_{n=1}^\infty$ as above, substituting integrals of μ for sums of μ. The assumption that μ is generating means in this context that the semigroup generated by the support of μ,

$$\mathrm{supp}(\mu) = \{g \in G\,;\, \mu(A) > 0 \text{ for any open subset } A \text{ containing } g\},$$

is the whole of G. This implies that $\{Y_n\}_{n=1}^\infty$ is irreducible. The proof that it is doubly stochastic and recurrent is identical to the proof above. □

Continuing the discussion of item (c) above, the return time τ of $\{Z_n\}_{n=1}^\infty$ to Γ is equal to the return time

$$\begin{array}{rcccl} \overline{\tau} & : & \Omega & \longrightarrow & \mathbb{N} \\ & & \omega & \longmapsto & \min\{n \geq 1\,;\, Y_n(\omega) = \Gamma e\} \end{array}$$

of $\{Y_n\}_{n=1}^\infty$ to the trivial coset Γe. In particular, $\mathbb{E}[\overline{\tau}] = \mathbb{E}[\tau]$ and thus $\{Y_n\}_{n=1}^\infty$ is a positive/null recurrent Markov chain if and only if Γ is a positive/null recurrent subgroup.

We are now ready to prove Theorems 1.2. In this proof we use some classical results on Markov chains, which the uninitiated reader may find in textbooks such as [8].

Proof of Theorem 1.2. Consider the definition of $\{Y_n\}_{n=1}^\infty$ above. Since $\{Y_n\}_{n=1}^\infty$ is irreducible and recurrent, it has a unique (up to multiplication by constants) stationary measure η. Since it is also doubly stochastic, this stationary measure is constant.

If $[G : \Gamma] < \infty$ then we can normalize η to be a probability measure, in which case, by Kac's theorem, $\mathbb{E}[\tau]$, the expected return time to Γe, is equal to $1/\eta(\Gamma e) = [G : \Gamma]$. If $[G : \Gamma] = \infty$ then η cannot be normalized to be a probability measure. Hence $\{Y_n\}_{n=1}^\infty$ admits no stationary probability measure, and is therefore null recurrent, so that $\mathbb{E}[\tau] = \infty$.

□

As a direct consequence of this Kac formula, the property of positive recurrence of a subgroup Γ of G is independent of μ: for any generating measure, the positive recurrent subgroups of G are those with finite index. On the other hand, the properties of null recurrence and transience do depend on the chosen measure μ. For example, let $G = \mathbb{Z}$ and $\Gamma = \{0\}$ be the trivial subgroup. Then Γ is recurrent for the simple random walk on $\mathbb{Z}$, but transient for any random walk on $\mathbb{Z}$ with drift. It may be interesting to characterize the subgroups that are transient for any generating measure μ on G.

4. An Abramov formula

In this section we prove Theorem 1.1. Let (X, ν) be a (G, μ)-stationary space. For shortness of notation, let h_μ and h_θ denote $h_\mu(X, \nu)$ and $h_\theta(X, \nu)$, respectively. Recall the definitions of the maps ρ and φ given by Eqs. 2.2 and 2.4:

$$\rho(g, x) = -\log \frac{dg^{-1}\nu}{d\nu}(x), \qquad \varphi(g) = \int_X \rho(g, x) d\nu(x).$$

Lemma 4.1. *For any $g \in G$,*

$$\sum_{g_1 \in G} \mu(g_1) \cdot \varphi(gg_1) = \varphi(g) + h_\mu\,. \tag{4.1}$$

Hence φ is, in a sense, a "nearly harmonic" function, in which the mean value property is corrected by a factor of h_μ.

Proof. For a fixed $g \in G$, integrate the cocycle relation

$$\rho(gg_1, x) = \rho(g, g_1 x) + \rho(g_1, x)$$

with respect to g_1 and x to get

$$\sum_{g_1 \in G} \mu(g_1) \cdot \varphi(gg_1) = \sum_{g_1 \in G} \mu(g_1) \cdot \int_X \rho(g, g_1 x) d\nu(x) + h_\mu\,.$$

Applying a change of coordinates and using the μ-stationarity of ν, the sum in the right hand side of the above equality is

$$\begin{aligned}\sum_{g_1 \in G} \mu(g_1) \cdot \int_X \rho(g, x) dg_1\nu(x) &= \int_X \rho(g, x) d\left(\sum_{g_1 \in G} \mu(g_1) \cdot g_1 \nu\right)(x) \\ &= \int_X \rho(g, x) d\nu(x) \\ &= \varphi(g),\end{aligned}$$

and the claim is proved. ∎

Proceeding as in the proof of Lemma 2.12, one may define the martingale $M_1, M_2, \dots$ by

$$M_n = \varphi(Z_n) - n \cdot h_\mu.$$

The expectation of each M_n is equal to $\mathbb{E}[M_0] = \varphi(e) = 0$. Note that, when this martingale or its increments are uniformly bounded, Theorem 1.1 is a consequence of the optional stopping theorem:

$$0 = \mathbb{E}[M_0] = \mathbb{E}[M_\tau] = \mathbb{E}[\varphi(Z_\tau) - \tau \cdot h_\mu] = h_\theta - \mathbb{E}[\tau] \cdot h_\mu.$$

However, this martingale is not bounded, and in general neither are its increments. For example, when μ has full support and $h_\mu \neq 0$, then $|M_n - M_{n+1}|$ can be arbitrarily large. Hence the optional stopping theorem cannot be used to prove Theorem 1.1. Instead, we replicate below a proof of the optional stopping theorem, keeping account of an error term R_n. To show that R_n vanishes, we prove in Appendix A a lemma for the Markov chain $\{Y_n\}_{n=1}^\infty$ defined in Section 3.

Let $t_n = \mathbb{P}[\tau = n]$ be the probability that the return time to Γ is equal to n. Start by writing Eq. 4.1 at $g = e$. Since $\varphi(e) = 0$, this becomes

$$h_\mu = \sum_{g_1 \in G} \mu(g_1) \cdot \varphi(g_1).$$

Separating the sum into values of g_1 that are in Γ and values that are not, we get

$$\begin{aligned} h_\mu &= \sum_{g_1 \in \Gamma} \mu(g_1) \cdot \varphi(g_1) + \sum_{g_1 \notin \Gamma} \mu(g_1) \cdot \varphi(g_1) \\ &= \theta^{(1)}(\varphi) + \sum_{g_1 \notin \Gamma} \mu(g_1) \cdot \varphi(g_1). \end{aligned}$$

To each term $\varphi(g_1)$ with $g_1 \notin \Gamma$ apply Eq. 4.1 again to obtain

$$\begin{aligned} h_\mu &= \theta^{(1)}(\varphi) + \sum_{g_1 \notin \Gamma} \mu(g_1) \left(\sum_{g_2 \in G} \mu(g_2) \cdot \varphi(g_1 g_2) - h_\mu \right) \\ &= \theta^{(1)}(\varphi) + \sum_{\substack{g_1 \notin \Gamma \\ g_2 \in G}} \mu(g_1) \cdot \mu(g_2) \cdot \varphi(g_1 g_2) - h_\mu \cdot \sum_{g_1 \notin \Gamma} \mu(g_1) \\ &= \theta^{(1)}(\varphi) + \sum_{\substack{g_1 \notin \Gamma \\ g_2 \in G}} \mu(g_1) \cdot \mu(g_2) \cdot \varphi(g_1 g_2) - h_\mu \cdot (1 - t_1) \\ &= \theta^{(1)}(\varphi) + \theta^{(2)}(\varphi) + \sum_{\substack{g_1 \notin \Gamma \\ g_1 g_2 \notin \Gamma}} \mu(g_1) \cdot \mu(g_2) \cdot \varphi(g_1 g_2) - h_\mu \cdot (1 - t_1). \end{aligned}$$

Then, using the fact that $\sum_{k \geq 1} t_k = 1$, we arrive at

$$h_\mu \cdot \left(\sum_{k \geq 1} t_k + \sum_{k \geq 2} t_k \right) = \theta^{(1)}(\varphi) + \theta^{(2)}(\varphi) + \sum_{\substack{g_1 \notin \Gamma \\ g_1 g_2 \notin \Gamma}} \mu(g_1) \cdot \mu(g_2) \cdot \varphi(g_1 g_2)$$

Recalling Definition 2.3 of the avoiding sets A_n, we can repeat the above argument to conclude that

$$h_\mu \cdot \sum_{j=1}^{n} \sum_{k \geq j} t_k = \sum_{k=1}^{n} \theta^{(k)}(\varphi) + \sum_{(g_1, \ldots, g_n) \in A_n} \mu(g_1) \cdots \mu(g_n) \cdot \varphi(g_1 \cdots g_n). \tag{4.2}$$

Observe that

$$\lim_{n \to \infty} \sum_{j=1}^{n} \sum_{k \geq j} t_k = \lim_{n \to \infty} \sum_{k=1}^{n} k \cdot t_k = \mathbb{E}[\tau],$$

which equals $[G : \Gamma]$, by Theorem 1.2. Observe also that

$$\lim_{n \to \infty} \sum_{k=1}^{n} \theta^{(k)}(\varphi) = \theta(\varphi) = h_\theta.$$

Hence, if we take the limit of Eq. 4.2 as n goes to infinity, we get

$$h_\mu \cdot [G : \Gamma] = h_\theta + \lim_{n \to \infty} \sum_{(g_1, \ldots, g_n) \in A_n} \mu(g_1) \cdots \mu(g_n) \cdot \varphi(g_1 \cdots g_n).$$

The theorem will be proved if we show that the error term

$$\begin{aligned} R_n &= \sum_{(g_1,\ldots,g_n)\in A_n} \mu(g_1)\cdots\mu(g_n)\cdot\varphi(g_1\cdots g_n) \\ &= \mathbb{E}\left[\varphi(Z_n)\cdot \mathbf{1}_{\{\tau>n\}}\right] \end{aligned}$$

converges to zero as n goes to infinity. To this purpose, we first bound $\varphi(g_1\cdots g_n)$.

Lemma 4.2. *For any $g_1,\ldots,g_n \in G$,*

$$\varphi(g_1\cdots g_n) \le -\sum_{k=1}^{n} \log\mu(g_k). \tag{4.3}$$

Proof. We first show that for any $g\in G$, $x\in X$ and $n\ge 1$ it holds that

$$\rho(g,x) \le -\log\mu^n(g). \tag{4.4}$$

This is stated in [7]. Note that, since ν is μ-stationary, it is also μ^n-stationary, that is, $\sum_{g\in G}\mu^n(g)\cdot g\nu = \nu$. Therefore, for any $x\in X$,

$$1 = \frac{d\nu}{d\nu}(x) = \frac{d\left(\sum_{g\in G}\mu^n(g)\cdot g\nu\right)}{d\nu}(x) = \sum_{g\in G}\mu^n(g)\cdot\frac{dg\nu}{d\nu}(x)$$

and since all the addends in the above sum are nonnegative, for any $g\in G$

$$\frac{dg^{-1}\nu}{d\nu}(x) \le \frac{1}{\mu^n(g^{-1})} \quad\Longrightarrow\quad \rho(g,x) \ge -\log\frac{1}{\mu^n(g^{-1})}\,.$$

Now, by the cocycle property (Eq. 2.3) we have

$$\rho(g,x) = -\rho(g^{-1},gx) \le -\log\mu^n(g),$$

thus establishing Eq. 4.4. Because this bound is independent of x, it implies

$$\varphi(g) = \int_X \rho(g,x)d\nu(x) \le -\log\mu^n(g).$$

To conclude the proof, observe that

$$\mu^n(g_1\cdots g_n) \ge \mu(g_1)\cdots\mu(g_n)$$

and so

$$\varphi(g_1\cdots g_n) \le -\log\mu^n(g_1\cdots g_n) \le -\sum_{k=1}^{n}\log\mu(g_k).$$

$\square$

Plugging this estimate in the error term R_n, we get

$$\begin{aligned} R_n &= \mathbb{E}\left[\varphi(Z_n)\cdot\mathbf{1}_{\{\tau>n\}}\right] \\ &\le \mathbb{E}\left[-\left(\sum_{k=1}^{n}\log\mu(X_k)\right)\cdot\mathbf{1}_{\{\tau>n\}}\right] \\ &= -\sum_{k=1}^{n}\mathbb{E}\left[\log\mu(X_k)\cdot\mathbf{1}_{\{\tau>n\}}\right]. \end{aligned}$$

Conditioning on X_k, we arrive at

$$R_n \le -\sum_{k=1}^{n}\sum_{g\in G}\mathbb{E}\left[\log\mu(X_k)\cdot\mathbf{1}_{\{\tau>n\}}|X_k=g\right]\cdot\mathbb{P}\left[X_k=g\right]$$
$$=-\sum_{k=1}^{n}\sum_{g\in G}\mu(g)\log\mu(g)\cdot\mathbb{P}\left[\tau>n|X_k=g\right],$$

since $\mathbb{P}[X_k=g]=\mu(g)$. Note that we are conditioning on the step X_k and not the position Z_k. By Lemma A.2, each of these conditional probabilities $\mathbb{P}[\tau>n|X_k=g]$ is bounded by e^{-Cn}, for some $C>0$ independent of n and k. Hence

$$R_n \le \sum_{k=1}^{n}\sum_{g\in G}-\mu(g)\cdot\log\mu(g)\cdot e^{-Cn}=n\cdot e^{-Cn}\cdot H(\mu),$$

which converges to zero as n goes to infinity. This concludes the proof of Theorem 1.1.

We finish this section by proving Corollary 1.3. Let (X,ν) be the Furstenberg-Poisson boundary of (G,μ). Corollary 2.14 guarantees that (X,ν) is also the Furstenberg-Poisson boundary of (Γ,θ) and thus, by Theorem 2.8,

$$h_\mu(X,\nu)=h(G,\mu)\quad\text{ and }\quad h_\theta(X,\nu)=h(\Gamma,\theta).$$

Corollary 1.3 thus follows by plugging the above equalities in Theorem 1.1.

5. Acknowledgements

The authors are thankful to The Weizmann Institute of Science for the excellent atmosphere during the preparation of this manuscript and to Uri Bader, Itai Benjamini, Hillel Furstenberg, Elchanan Mossel and Omri Sarig for valuable comments and suggestions.

Appendix A. A lemma on the Markov chain $\{Y_n\}_{n=1}^\infty$

Let $\{M_n\}_{n=1}^\infty$ be an irreducible time-independent Markov chain in the finite state space S. For each $x\in S$, let

$$\tau_x=\min\{n\ge 1\,;\,M_n=x\}$$

be the hitting time for x and let $\mathrm{Prob}_x[\,\cdot\,]$ denote the probability in $\{M_n\}_{n=1}^\infty$ given that $M_0=x$. The result below is standard in the theory of Markov chains (see, e.g., the chapter on finite Markov chains in [8]).

Lemma A.1. *Under the above condition, there exists $C>0$ such that*

$$\mathrm{Prob}_x\left[\tau_y>n\right]\le e^{-Cn}$$

for all $x,y\in S$ and $n\ge 1$.

The lemma used to bound the error term R_n in the proof of Theorem 1.1 can now be proved. We henceforth assume the conditions and use the notation of Sections 3 and 4.

Lemma A.2. *There exists $C>0$ such that*

$$\mathbb{P}\left[\tau>n\,|\,X_k=g\right]\le e^{-Cn}$$

for all $n,k\ge 1$ and $g\in G$.

Proof. Apply the previous lemma to the irreducible time-independent Markov chain $\{Y_n\}_{n=1}^{\infty}$ in the finite set $\Gamma\backslash G$ as in Section 3, defined by the projection of the Markov chain $\{Z_n\}_{n=1}^{\infty}$ in G, to get $C > 0$ such that

$$\mathbb{P}_g\left[\tau > n\right] \leq e^{-2Cn}$$

for $n \geq 1$ and $g \in G$, where $\mathbb{P}_g$ denotes the measure of μ-random walks starting at g. We divide the proof into two cases.

Case 1: $k > n/2$. Because the event $\{\tau > n/2\}$ is independent of X_k, we have

$$\begin{aligned}\mathbb{P}\left[\tau > n \,|\, X_k = g\right] &\leq \mathbb{P}\left[\tau > n/2 \,|\, X_k = g\right] \\ &= \mathbb{P}\left[\tau > n/2\right] \\ &\leq e^{-Cn}.\end{aligned}$$

Case 2: $k \leq n/2$. By the law of conditional probabilities, $\mathbb{P}\left[\tau > n \,|\, X_k = g\right]$ is equal to

$$\sum_{\Gamma g_1 \in \Gamma\backslash G} \mathbb{P}\left[\tau > n \,|\, Y_{n/2} = \Gamma g_1, X_k = g\right] \cdot \mathbb{P}\left[Y_{n/2} = \Gamma g_1 \,|\, X_k = g\right].$$

We condition the first term on $\tau > n/2$, to get

$$\begin{aligned}\mathbb{P}\left[\tau > n \,|\, Y_{n/2} = \Gamma g_1, X_k = g\right] &\leq \mathbb{P}\left[\tau > n \,|\, \tau > n/2, Y_{n/2} = \Gamma g_1, X_k = g\right] \\ &= \mathbb{P}\left[\tau > n \,|\, \tau > n/2, Y_{n/2} = \Gamma g_1\right] \\ &= \mathbb{P}_{g_1}\left[\tau > n/2\right],\end{aligned}$$

where the first equality follows from the Markov property. Hence

$$\begin{aligned}\mathbb{P}\left[\tau > n \,|\, X_k = g\right] &\leq \sum_{\Gamma g_1 \in \Gamma\backslash G} \mathbb{P}_{g_1}\left[\tau > n/2\right] \cdot \mathbb{P}\left[Y_{n/2} = \Gamma g_1 \,|\, X_k = g\right] \\ &\leq e^{-Cn} \sum_{\Gamma g_1 \in \Gamma\backslash G} \mathbb{P}\left[Y_{n/2} = \Gamma g_1 \,|\, X_k = g\right] \\ &= e^{-Cn},\end{aligned}$$

thus completing the proof of the lemma. □

References

[1] L.M. Abramov, *The entropy of a derived automorphism (Russian)*, Dokl. Akad. Nauk SSSR **128** (1959), 647–650.

[2] A. Avez, *Entropie des groupes de type fini (French)*, C. R. Acad. Sci. Paris Sér. A-B **275** (1972), 1363–1366.

[3] U. Bader and Y. Shalom, *Factor and normal subgroup theorems for lattices in products of groups*, Invent. Math. **163** (2006), no. 2, 415–454.

[4] A. Furman, *Random walks on groups and random transformations*, Handbook of dynamical systems **1A** (2002), 931–1014.

[5] H. Furstenberg, *Random walks and discrete subgroups of Lie groups*, Advances in Probability and Related Topics **1** (1971), 1–63.

[6] Hillel Furstenberg and Eli Glasner, *Stationary dynamical systems*, Dynamical numbers—interplay between dynamical systems and number theory, 2010, pp. 1–28. MR2762131

[7] V.A. Kaimanovich and A.M. Vershik, *Random walks on discrete groups: boundary and entropy*, Ann. Probab. **11** (1983), no. 3, 457–490.

[8] D.A. Levin, Y. Peres, and E.L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009.

[9] A. Nevo and R.J. Zimmer, *Homogenous projective factors for actions of semi-simple Lie groups*, Invent. Math. **138** (1999), no. 2, 229–252.

Weizmann Institute of Science, Faculty of Mathematics and Computer Science, POB 26, 76100, Rehovot, Israel.

E-mail address: `yair.hartman, yuri.lima, omer.tamuz@weizmann.ac.il`